Evolutionary Applications
Open Access
Evolutionary approaches to environmental, biomedical and socio-economic issues

# Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow



| | |
|---|---|
| Journal: | *Evolutionary Applications* |
| Manuscript ID | EVA-2018-272-SOA.R3 |
| Manuscript Type: | Special Issue Original Article |
| Keywords: | Landscape genomics, Conservation Genetics, environmental association analysis, adaptive potential, assisted gene flow, provenance matching |
| | |

# Abstract

When populations of a rare species are small, isolated and declining under climate change, some populations may become locally maladapted. Detecting this maladaptation may allow effective rapid conservation interventions, even if based on incomplete knowledge. Population maladaptation may be estimated by finding genome-environment associations (GEA) between allele frequencies and environmental variables across a local species range, and identifying populations whose allele frequencies do not fit with these trends. We can then design assisted gene flow strategies for maladapted populations, to adjust their allele frequencies, entailing lower levels of intervention than with undirected conservation action. Here, we investigate this strategy in Scottish populations of the montane plant dwarf birch (*Betula nana*). In genome-wide restriction site-associated single nucleotide polymorphism (SNP) data we found 267 significant associations between SNP loci and environmental variables. We ranked populations by maladaptation estimated using allele frequency deviation from the general trends at these loci; this gave a different prioritization for conservation action than the Shapely Index, which seeks to preserve rare neutral variation. Populations estimated to be maladapted in their allele frequencies at loci associated with annual mean temperature were found to have reduced catkin production. Using an environmental niche modelling (ENM) approach, we found annual mean temperature (35%), and mean diurnal range (15%), to be important predictors of the dwarf birch distribution. Intriguingly, there was a significant correlation between the number of loci associated with each environmental variable in the GEA, and the importance of that variable in the ENM. Together, these results suggest that the same environmental variables determine both adaptive genetic variation and species range in Scottish dwarf birch. We suggest an assisted gene flow strategy that aims to maximize the local adaptation of dwarf birch populations under climate change by matching allele frequencies to current and future environments.

## Keywords

Landscape genomics, conservation genetics, environmental association analysis, evolutionary conservation, adaptive potential, climate change, assisted gene flow, provenance matching.

## Introduction

Climate change is predicted to become a major driver of global biodiversity loss (Bellard et al., 2012; Urban, 2015). Species that lack relevant phenotypic plasticity (Gratani, 2014; Nicotra et al., 2010) may survive environmental changes by dispersing to new locations, consequently tracking conditions they are currently adapted to (Aitken et al. 2008; Meier et al. 2012), or remaining in the same location and rapidly evolving adaptation to their new environments from standing genetic variation or gene flow (Aitken et al., 2008; Alberto et al., 2013). Migration in response to rapid climate change may be particularly difficult for plants (Corlett and Westcott, 2013; Hampe and Petit, 2005; Zhu et al., 2012). In some cases plants lack the dispersal ability to keep pace with accelerated climate shifts (Loarie et al., 2009). For example, there may be an absence of potential habitat at higher latitudes (McKenney et al., 2007) or altitudes (Engler et al., 2011), suitable new habitats may be separated by too large distances (Meier et al., 2012) or dispersal may be impossible due to anthropogenic habitat fragmentation. In these cases, conservation managers aiming to prevent extinction of species or populations face a choice between relying on *in situ* evolution to track the environmental change, or attempting conservation interventions such as assisted migration or assisted gene flow that seeks to enable, facilitate or accelerate adaptation.

To evaluate whether interventions are appropriate, a first step is understanding current local adaptation and the potential for adaptation to future environments (Davis et al., 2005; Hoffmann et al., 2017; Funk et al 2019). The classical way to identify local adaptation is via reciprocal transplant experiments (Kawecki and Ebert, 2004; Leimu and Fischer, 2008; Pardo-

Diaz et al., 2015). However, this approach is often unfeasible for wild organisms with long generation times in need of urgent conservation, meaning that more rapid approaches using genomics are desirable (Williams et al., 2008).

Genotype-environment association (GEA; also referred to as environmental association analysis, EAA) methods are increasingly used to identify loci involved in local adaptation (Abebe et al., 2015; Ahrens et al., 2018; Bay et al., 2017; Coop et al., 2010; Flanagan et al., 2018; Günther and Coop, 2013; Rellstab et al., 2015; Funk et al 2019). These approaches detect replicated signatures of selection (SNPs that deviate strongly from estimated neutral population structure) across many independent populations. Thus far the majority of studies to apply GEA in tree species have been targeted at candidate genes, and surveyed fewer than 350 loci (Keller et al., 2012; Nadeau et al., 2016; Rellstab et al., 2016; Wang et al., 2016).

Building on the assumption that GEA captures an important component of locally adaptive allelic variation, especially if based on genome-wide markers, we may extend it to rapidly assess local adaptation and adaptive potential within populations. The principal behind this approach is the detection of discordance between genotype and environment, in certain populations, as an indicator of reduced local adaptation and vulnerability to future demographic decline (Alberto et al., 2013). In a previous study, Rellstab *et al.* (2016) developed a model to estimate the average change in allele frequency at environmentally-associated loci that would be required to respond to projected future environmental conditions. They based this estimate on the allele frequency changes that would maintain the present-day associations between genotype and environment and termed this mismatch, the risk of non-adaptedness (RONA). For clarity we term this 'future risk of non-adaptedness' (f-RONA) and comment that rather than a 'risk' this is a forecast, but for consistency we maintain the same terminology in this manuscript. This approach to estimating adaptation has many simplifying assumptions. Environmental variation in nature is complex, as are the mechanisms by which organisms adapt to them, but as Funk et al (2019) argue, any available evidence may improve conservation decision making.

Here, we extend the work of Rellstab *et al.* (2016) to explicitly define c-RONA, the 'current risk of non-adaptedness', that is the average change in allele frequency at climate-associated loci required to match our estimate of the optimum for current climatic conditions (for a given environmental factor). Current risks are likely to be particularly important for species that are already declining due to climate change, and have small isolated populations. Furthermore, we extend the univariate RONA model to a multi-locus analysis of genome-wide markers, and use best linear unbiased prediction (BLUP) to improve our estimate of the effect of each locus.

In populations where c-RONA is high, local genotypes would not match local environmental variables as expected. Therefore, a possible management intervention is to use assisted gene flow (AGF) to introduce more appropriate alleles or adjust population allele frequencies. Here, AGF is defined as the managed movement of individuals or gametes between populations, from source populations that have been selected with the aim of accelerating adaptation, so that it is faster than would occur by passive natural dispersal alone (Aitken and Whitlock, 2013). This AGF strategy could be used to inform sourcing of seed stock for reforestation programs (Boshier et al., 2015) and mitigate maladaptation to future climate (Aitken and Bemmels, 2016; Havens et al., 2015; Jin et al., 2016). Importantly, only modest translocation of genotypes may enhance adaptation by introducing genetic variation upon which selection can act to further refine local allele frequencies (Bay et al., 2017; Pavlova et al., 2017). Conversely such interventions could have negative effects (i.e. outbreeding depression) if they cause gene flow between populations with undetected adaptive differentiation (Frankham et al., 2011; Pavlova et al., 2017). We note that where target populations are small, maladapted and dominated by drift, Assisted Gene Flow is equivalent to Genetic Rescue (see Aitken and Whitlock (2013) for a detailed review).

If AGF is to be effective, there must be appropriate populations from which to source migrants. Such populations might be found towards the species' retreating range edge or other locations where environmental conditions are closer to those anticipated in the future (Olson et al.,

2013). To design a sampling strategy that encompasses both environmental gradients and declining range edge populations threatened by environmental change, we can use environmental niche models (ENMs) (Maguire et al., 2015). ENMs project the distribution of species' ranges under current and future climate scenarios based on observation data and can guide effective sampling (Elith and Leathwick, 2009). ENMs are also an established tool for conservation practitioners seeking to understand major climatic selection pressures and projected range shifts for threatened species, but often lack integration and comparison with genomic assays of local adaptation (Hällfors et al., 2015; Razgour et al., 2019).

Here, we conduct GEA and ENM analysis of wild populations of dwarf birch (*Betula nana*), for which we have field observation and genome-wide population genetic data. In the UK, dwarf birch is a nationally scarce montane tree that has experienced an accelerated decline in recent decades, likely due to the combined impact of anthropogenic climate change and moorland management that permits over-browsing and burning (Aston, 1984; Borrell et al., 2018; Wang et al., 2014; Zohren et al., 2016). Dwarf birch, like many tree species, is the focus of a conservation program to restore populations, delimit management units and prioritise the protection of important genetic diversity (Koskela et al., 2013). Germplasm collection from central Scottish Highland populations is already underway for reintroduction to other parts of the species former range (pers. obs. J Borrell). Previous research by our group has found that despite extensive fragmentation, most populations of dwarf birch in the UK contain diversity comparable to that of large, unfragmented Scandinavian populations (Borrell et al., 2018). Nevertheless, we concluded that this diversity has become increasingly partitioned among populations. In other words, much of the adaptive diversity in dwarf birch is still extant in the UK, but due to restricted gene flow and dispersal, marginal populations may be maladapted due to a failure to track environmental change, or by drift of adaptive alleles away from their optimum frequency. There is limited potential for naturally occurring gene flow to enhance future adaptation in many populations.

In species subject to conservation management such as dwarf birch, evolutionary processes have sometimes been overlooked, despite the importance of adaptation to species persistence (Eizaguirre and Baltazar-Soares, 2014; Fitzpatrick and Keller, 2015). Therefore the adaptive potential of populations may be underrepresented in conservation prioritization strategies (Funk et al., 2019; Harrisson et al., 2014). For example, where genetic diversity information is available to conservationists, metrics that score populations on neutral genetic distinctiveness, such as the Shapley Index are often used (Haake et al., 2007; Isaac et al., 2007; Volkmann et al., 2014). However there is no guarantee that neutral and adaptive diversity will be correlated (Bonin et al., 2007), and indeed approaches designed solely to promote or conserve neutral diversity may be harmful (Reed and Frankham, 2003; Weeks et al., 2016). Therefore evaluating adaptive diversity, rather than using more established metrics of genetic diversity should improve the prioritisation decisions in species management, though see Kardos and Shafer, (2018) for potential pitfalls.

To explore potential management strategies for dwarf birch, that takes into account local adaptation and evolutionary potential, we first characterise the species' range using ENMs under present and projected future climate scenarios. We evaluate these ENMs by assessing whether populations on the margins of the inferred distribution had lower scores for phenotypic and fitness proxies for local adaptation. Second, we use GEA to survey putative adaptive loci across the species' range and estimate c-RONA to identify populations with a discordance between genotype and environment. The combined ENM and GEA data present an opportunity to test the hypothesis that limiting environmental variables (which have higher discriminatory power in an ENM) have more genomic loci associated with them in GEA, perhaps as a result of stronger selection for adaptation (an alternative would be that certain variables limit species' ranges precisely because they lack genetic adaptation). We provide preliminary evidence in support of this hypothesis in dwarf birch. Third, we evaluate our estimates of non-adaptedness (c-RONA) of dwarf birch populations against the Shapley Index, an existing

conservation prioritization most often applied to neutral markers. Finally, we illustrate a strategy of AGF to maximize adaptive genetic diversity and hence sustain the adaptive potential of British dwarf birch populations. We discuss the advantages and limitations of this approach in the context of managing dwarf birch and other plants exposed to rapid environmental change.

## Methods

### Environmental niche modelling

To determine the environmental variables influencing the present and future distribution of dwarf birch in the UK, we developed an ENM based on 763 resampled fine-scale (≤1 km) records from the period 1960-present. Records were sourced from national databases, conservation partners and fieldwork observations (see Borrell *et al.* 2018). Nineteen bioclimatic layers were obtained from the WorldClim database (www.worldclim.org) at 1km resolution (Hijmans et al., 2005), for the period 1960-1990, including 11 temperature and eight precipitation derived variables reflecting annual trends, seasonality and limiting environmental factors. High resolution elevation data was used to compute slope and aspect terrain characteristics using the *Raster* package (Hijmans & Etten, 2012) in R software (R Development Core Team, 2014). These variables are indicators of soil moisture, erosion, wind and solar radiation (Hoersch et al., 2002). To avoid overfitting, we removed multiple highly correlated variables (correlation coefficient >0.7), retaining 10 for analysis (preferring less derived, e.g. Annual Mean Temperature, rather than Monthly or Quarterly values) (Table 1, Figure S1). Elevation was excluded due to its high correlation with temperature (Parolo et al., 2008). Temperature was retained because it captures the projected change in climate change models, whilst elevation does not. All retained variables were standardized to a mean of zero and unit variance. Eight further datasets consisting of the same retained variables were generated under four representative concentration pathways (RCP) defined by the

Intergovernmental Panel on Climate Change Fifth Assessment (IPCC, 2014a) at each of two future time points (2045-65 and 2081-2100). These projections allow estimation of future temperature and precipitation values across the study area derived from the Community Climate System Model (Gent et al., 2011) (Table S1).

The ENMs were generated using MaxEnt (Phillips et al., 2006) within the *dismo* package (Hijmans et al., 2011). We performed 50 randomly subsampled replicate runs with 25% of observations retained for cross-validation. Models were further evaluated using a binomial test of omission rate and Area Under the Receiver Operating Characteristic Curve (AUC). A species occurrence threshold to assess changes in occupied area was defined by 'maximum training sensitivity plus specificity', which optimizes the trade-off between commission and omission errors (Liu et al., 2016). Rank and percentage contribution of environmental variables is reported here, as these have been demonstrated to capture biologically important factors (Searcy & Shaffer, 2016).

**Phenotypic data and habitat suitability projections**

We identified 29 dwarf birch populations that encompass the extant UK range (Table 2, Figure S2). To test the performance of our ENM, we collected extensive phenotypic measurements of traits related to reproductive output and fitness in 20-30 individuals per population in June-August 2013. These included: the number of male and female catkins, plant area, plant height and diameter of the largest stem. Cambial tissue samples were retained for genetic analysis. A subset of 18 populations was also tested for seed viability in germination experiments, a fitness proxy relevant to population persistence (Alsos et al., 2003). Seed were collected in late summer, over-wintered at 4°C then kept in moist conditions at 18-20°C with a 14h photoperiod for 60 days the following spring. For nine of these populations, 100-day survival of seedlings during the following Spring was measured (See Supplementary Materials for details).

To assess change in habitat quality across the study area, we first plotted the ENM derived habitat suitability index (HSI) estimates for all populations under current and future conditions. Second, ENM performance was assessed using a generalized linear model with a quasipoisson error distribution to test for a relationship between present time HSI estimates and mean population catkin counts. We also tested for a relationship between HSI (explanatory variable) and mean germination rates (response variable) using a quasibinomial error distribution. Here we are explicitly testing the hypothesis that plants displayed greater reproductive output in locations with a higher ENM derived HSI.

**RAD sequencing**

The genetic samples used in this study are a subset of those described in (Borrell et al., 2018). Briefly, DNA was extracted from 130 individuals (Table 2) and submitted to Floragenex (Oregon, USA) for 100bp single-end RAD sequencing with the enzyme *PstI*. Raw reads were filtered using Stacks v1.35 (Catchen et al., 2013) and aligned to the dwarf birch genome, retaining only reads that align uniquely (Wang et al., 2013) using Bowtie2 (Langmead and Salzberg, 2012) and the *ref_map.p*l pipeline. SNPs were called with a minimum depth of 5, the bounded model and a minimum log likelihood of -20, with corrections made using *rxstacks*. Finally, we filtered for loci present in ≥8 populations, and a minor allele frequency >0.05.

**Genomic signatures of local adaptation**

We first used BayeScan (Foll and Gaggiotti, 2008) to compare allele frequency differences among populations and identify $F_{ST}$ outlier loci, so that these could be excluded for generating a null covariance matrix for Bayenv2. Analysis was performed with 50,000 iterations thinned every 10, with 20 pilot runs, a burn-in of 50,000 iterations and other parameters at default. Whilst $F_{ST}$ outliers are candidate loci of adaptation, they can also emerge because of selection due to deleterious alleles, hybrid zones and historical demography (Bierne et al., 2013).

Thusrelaxed BayeScan parameters allowed us to screen outlier loci prior to GEA analysis in Bayenv2 (Günther and Coop, 2013).

Bayenv2 incorporates neutral genetic structure using a covariance matrix based on neutral markers and attempts to identify correlations between outliers and environmental gradients, potentially reducing false positives (De Mita et al., 2013). Based on recommendations in François *et al.* (2016), to further minimize false positives we initially excluded loci detected in BayeScan to compute a null covariance matrix of relatedness between populations, over 100,000 iterations and five independent runs. We then tested all loci (including those initially identified by BayeScan) under an alternative model where allele frequencies are determined by a combination of the covariance matrix and an environmental variable. We performed our analysis independently across all environmental variables, with the expectation that correlated predictors would return subsets of the same markers. The posterior probability that a locus is under selection, across each independent environmental variable was assessed using Bayes factors (BF), with log10 posterior odds ratio values >1 defined as strong support (Jeffery, 1961). We averaged BFs over independent runs as recommended by Blair et al. (2014), and following Günther & Coop (2013) we retained loci as good candidates if, in addition to a high BF, they also fell in the top 10% of Spearman correlation coefficient values, to further reduce false positives. For comparison, we also independently tested for signatures of local adaptation using Redundancy Analysis (RDA) (Forester et al., 2018; Rellstab et al., 2015), (see Supplementary Materials) though we consider only the candidates identified using Bayenv2 in subsequent analyses.

**Neutral and adaptive population structure**

To evaluate population structure, pairwise population $F_{ST}$ was computed in Arlequin v3.5.2 (Excoffier and Lischer, 2010), and performed separately for putative neutral and adaptive loci identified through GEA analysis using a method similar to that of Candy et al. (2015).

**Gene expression**

To provide an additional line of evidence on the activity of our candidate adaptive loci, we extracted up to 10,000bp flanking each side of the candidate locus from the *B. nana* reference genome and searched for these sequences in an RNA expression database using dwarf birch tissues derived from our genome reference plant under glasshouse conditions (Wang et al., 2013). Briefly, RNA was extracted from fresh dwarf birch leaves and flowers using a modified RNAeasy Plant Mini Kit (Qiagen, Hilden, Germany), incorporating additional CTAB and phenol-chloroform steps to generate 100bp paired-end reads with an average insert size of 280bp (for full methods see Zohren, 2016). These were mapped to the reference genome using Trinity software (Grabherr et al., 2013).

**Maladaptation under present and future conditions**

We carried out RONA analysis on the nine standardized environmental variables that were associated with six or more candidate loci, allocating each locus to the single environmental variable with the largest Bayes factor (thereby avoiding double-counting a locus in the c/f-RONA calculations below). We estimated the vector of effect sizes, $\boldsymbol{\beta}$, in which each value corresponds to a locus, using R package rrBLUP (Endelman 2011). In this analysis, the vector of allele frequencies $\boldsymbol{f}$ for each population was used as the predictor of the environment in that location. The sum of $\boldsymbol{f\beta}$ gives an estimate of the environment (the value of the environmental variable) to which the population would be best adapted. The residual deviation of the observed value from this expectation is a measure of the deviation from the optimum environment for that population (c-RONA), and is proportional to the change in allele frequency that would be required to match the population to its local environment (weighted by $\boldsymbol{\beta}$). This measure is therefore analogous to those employed by Rellstab *et al.* (2016) and Pina-Martins *et al.* (2018), which quantify the mismatch between genotypes and environment in terms of allele

frequencies. We combined information across variables by calculating the mean of the absolute residuals. Similarly, we could calculate the difference from the projected values of the environmental values under each climate change scenario to estimate f-RONA (Figure 2).

**Conservation prioritization**

We compared the magnitude of c-RONA across dwarf birch populations with the Shapley index (Haake et al., 2007). The Shapley index prioritizes populations based on evolutionary isolation and contribution to overall diversity based on pairwise differentiation. Several similar metrics are widely used for conservation management (Collen et al., 2011; Gumbs et al., 2018; Jetz et al., 2014). Here, we used the method outlined in Volkmann et al. (2014), which maximizes within-species genetic diversity using a network approach implemented in NeighborNet (Bryant and Moulton, 2004; Huson and Bryant, 2006). We used linear regression to test for a relationship between absolute c-RONA values and the Shapley index for neutral and adaptive loci.

**Simulated assisted gene flow**

For each environmental variable, and for each population in the study, we identified the population most appropriate for AGF based on the match between the local environment and the sum of $f\beta$. Where several suitable populations were identified within the confidence interval of our regression, we selected the location geographically closest to the recipient population, since there could be local adaptation to undetected environmental variables (cf. Boshier *et al.* 2015).

**Method validation and ENM-GEA comparison**

To validate our model we tested the hypothesis that higher c-RONA values would be associated with the reduced performance of fitness proxies. Therefore we tested for a correlation between population c-RONA values for each environmental variable, or their interactions, and i) square

root transformed catkin counts and ii) germination rate across study populations. Finally, we tested for a correlation between the relative importance of environmental variables identified in our ENM and the number of GEA loci associated with each variable.

# Results

## Environmental niche models

The dwarf birch ENM was well parameterized with high mean test AUC (0.946 ±0.008) and a low mean test omission rate (0.09, $p<0.001$) at a logistic threshold of occurrence of 0.193. Four variables together contributed >85% to the predictive model performance including annual mean temperature (34.9%) and maximum temperature of the warmest month (22.1%) (Table 1). The resulting model is highly concordant with qualitative field observations and inspection of variable curves showed biologically plausible responses (Figure S3). Future projections show significant declines across the species' range with persistent populations restricted to areas of higher elevation (Figures 1, S4). Excluding other anthropogenic pressures, under the most severe scenario (RCP8.5, 2081-2100), suitable habitat may be reduced to ~1% of the current extent (Table S2).

## Phenotypic data and habitat suitability

Phenotypic data means are reported in Table S3. Germination success was assayed in 190 individuals, and averaged 7.6% for both years with 6.1% 100-day survival (i.e. 80% of those that germinated) with substantial variation among populations (Table S4). A single large outlier individual (Emblehope) produced an exceptionally large number of catkins strongly biasing results, thus was excluded from subsequent analysis. Present time habitat suitability index (HSI) estimates for dwarf birch ranged from 0.0006 to 0.81 (Table 2), with substantial declines under all future scenarios (Figure S4). We found a significant non-linear positive relationship

between HSI and mean population catkin count ($F_{1,26}$=7.50, P=0.011) as well as HSI and the proportion of seeds that germinated ($F_{1,16}$=9.52, P=0.007) (Figure 1).

**RAD Sequencing and genotype-environment associations**

After quality control, RAD sequencing produced 173,460,998 reads, of which 79.1% aligned to the *B. nana* genome. Subsequently 73.2% of aligned reads mapped to a single unique position. Three samples were excluded due to low coverage. After filtering we retained 14,889 SNPs over 8,727 contigs. These contigs together cover approximately a third of the dwarf birch genome assembly. Bayescan identified 382 putative outlier SNPs at a conservative false discovery rate of 0.2, meant that we were more likely to remove false positives than leave false negatives. These were excluded during the generation of the Bayenv2 null covariance matrix. Subsequent GEA analysis detected 267 highly significant locus-environment associations, encompassing 303 SNPs (Table S5), with a single SNP from each locus retained for subsequent analysis. The most frequent associations were between mean diurnal range and 71 loci, and annual mean temperature and 64 loci, whereas variables such as temperature seasonality and mean temperature of driest or wettest quarters had comparatively few associated loci. Just six loci were in common between Bayescan and Bayenv2 detection methods, and Bayescan candidate loci did not report significantly higher BF scores compared to the dataset as a whole. A comparison between bayenv2 and RDA found highly significant correlation (Pearson's r(6) = 0.84, p = 0.008) between methods, in the number of genotype-environment associations identified for each environmental variable (Table S6, Figure S5) suggesting that both methods are identifying a similar genomic pattern of adaptation.

**Neutral and adaptive population structure**

Pairwise $F_{ST}$ values between populations ranged from 0.000 to 0.701 for putative neutral markers (mean = 0.100, n=14,889) and 0.000 to 0.260 for putative adaptive markers (mean = 0.079, n=303). We found more significant pairwise $F_{ST}$ values for adaptive markers (92 of 312

pairwise comparisons) than for neutral markers (49 of 312 pairwise comparisons) (Table S7). We note particularly that neutral pairwise $F_{ST}$ was upwardly biased by very small range edge populations (LX, EM, SA, BG and TD). If these populations are excluded mean pairwise neutral $F_{ST}$ is 0.069 and mean pairwise adaptive $F_{ST}$ is 0.076.

**Expression of putative adaptive loci**

The 267 loci mapped to 185 unique scaffolds in our reference genome. Based on RNAseq data, 35 candidate regions showed evidence of gene expression in flower tissue (19%), 15 showed gene expression in leaf tissue (8%) and 13 showed gene expression in both (7%). In comparison to the overall SNP dataset, we found that both flower ($X^2$=23.14, p<0.001) and leaf ($X^2$=8.59, p=0.003) expressed sequences are significantly over-represented among putatively adaptive loci.

**Potential for adaptation and conservation prioritization**

Mean population c-RONA based on environmentally associated SNPs under present climate was 0.22 (±0.10), ranging from from 0.07 (SE ±0.06) at Glen Cannich, to 0.39 (±0.24) at Beinn Enaiglair on the Western periphery of the species range (Table 2, S8). BLUP estimates for all variables are presented in Figure S6. Under future climate scenarios mean population f-RONA was greater than c-RONA, which increased from 0.22 (±0.10) to a maximum of 0.27 (±0.11) under scenario RCP8.5 (Table S9), with substantial variation across populations and projections. We found positive correlation between c-RONA and the Shapley Index for neutral genetic diversity (Pearson's r(24) = 0.44, p=0.023), despite a number of outliers as shown by the low correlation coefficient, but no such pattern for putative adaptive genetic diversity (Pearson's r(24) = 0.004, p=0.983) (Figure 3). The Shapley Index for neutral diversity also strongly favoured a small number of relict and range edge populations dominated by drift (e.g. BG, SA, see Borrell et al., 2018) whereas for adaptive diversity, the range of values was narrower suggested more even support across populations. Therefore, the Shapley Index and our metric

for maladaptation (c-RONA) provide very different ranking for conservation value (Table 2). A consensus ranking of populations is provided in Table S10.

### Simulating assisted gene flow

For each population across each environmental variable we identified the geographically closest 'donor' population with an allele frequency that would reduce c-RONA (within confidence limits) at the 'recipient' site (Figure 4, S7). This strategy proposes a pattern of dispersal from the centre of the distribution towards the periphery, particularly at the Southern range edge, though there are exceptions such as transfer from the Northern to Southern range edge (e.g. MTColdQ, Figure S7). In some cases, the analysis does not indicate the need for AGF in particular populations, such as those at the centre of the species distribution which appear to be well matched to their environment (i.e. locally adapted).

### Method validation and ENM-GEA comparison

If c-RONA values do indeed quantify the degree of maladaptation, they should be negatively correlated with independent measurements of population fitness. The c-RONA values for annual mean temperature (AMTemp) were significantly negatively correlated with mean population catkin counts ($F_{1,23}$=5.84, p=0.025) (Figure 5A) (we found a similar relationship for c-RONA averaged across all environmental variables, data not shown). The interaction of c-RONA for Annual Mean Temperature and Mean Diurnal Range correlated with germination rate ($F_{11,14}$=8.07 , p=0.004). Finally, in a comparison of ENM and GEA methods, we found a significant correlation between the number of genotype-environment associations and the percentage contribution of environmental variables defining species range in our ENM (r(8) = 0.69, p=0.027) (Figure 5B).

## Discussion

Environmental niche modelling projects that the decline of dwarf birch across the UK is likely to continue and become increasingly severe, with almost total range loss possible by the end of the century under the highest emission scenarios. We found that catkin production and seed germination are positively correlated with ENM projections of habitat suitability. This suggests lower reproductive fitness of plants in populations with lower habitat suitability index. We cannot fully exclude the possibility that low seed germination rates are partly due to high dormancy, but it is not obvious in this context that dormancy would increase fitness. Temperature was particularly important to our ENM projections, and previous work has shown reduced production of germinable seeds by dwarf birch in warmer climates (Alsos *et al.* 2003). In future, an overall decline in habitat suitability across the species' British range is likely to further reduce reproductive fitness and subsequent population persistence.

Genome-wide analysis identified 267 significant genotype-environment associations (0.018 of loci surveyed) across 24 environmental variables, which is consistent with the number of associations identified in similar studies (Abebe *et al.* 2015; Manthey and Moyle 2015; reviewed in Ahrens *et al.* 2018). These loci were significantly more commonly found within 10kb of a gene annotated on our reference genome sequence with cDNA evidence for expression than were SNP loci that were not identified as candidates, increasing our confidence that candidate loci could be involved in phenotypic traits.

We observe that of the four environmental variables that contribute substantially to the dwarf birch ENM (Table 1) three of these also account for the largest number of associated loci in the genotype-environment analysis (GEA) (Table 1, Table S5). Therefore, in a comparison of the two methods, we find significant agreement between ENM and GEA results in identifying important environmental variables (Figure 5B). It is not a logical necessity for environmental variables with the largest effects on species range limits to show the strongest correlation with allele frequencies. However, it is an interesting finding that suggests that we have identified biologically relevant environmental variables that influence both distribution and local

adaptation of dwarf birch. It would be valuable to test for this pattern in other species, in the context of genetic models of species range limits (Polechová, 2018; Polechová and Barton, 2015).

We surveyed the allele frequencies of these GEA loci across populations to estimate c-RONA. As expected, we find the populations which we have identified as having a poor match between genotype and environment (high c-RONA) are particularly small or isolated, and those on the margins of the species' distribution. This result is consistent with reconstruction of demographic history and genetic differentiation by Borrell *et al.* (2018), where we showed that populations with a census size of less than 10 (LX, EM, SA, BG and TD) had unusually high levels of $F_{ST}$. In this previous study we estimated the maximum likelihood value of local $F_{ST}$ relative to the regional mean, using the multinomial Dirichlet likelihood function proposed by Balding and Nichols (1995) and evaluated the influence of sample size by estimating ML-$F_{ST}$ across all loci from a single individual drawn from each population (Borrell et al (2018)). From this we concluded that these small populations were suffering from severe genetic drift. Mean pairwise $F_{ST}$ for these small populations is 0.331 for neutral markers and 0.116 for putative adaptive markers. Whereas for the remaining 21 populations mean pairwise $F_{ST}$ is 0.069 for neutral markers and 0.076 for adaptive markers. This suggests that in healthy populations there is more differentiation at loci under selection, as expected. We also found that c-RONA estimates for annual mean temperature were negatively correlated with mean population catkin counts and the interaction of c-RONA for Annual Mean Temperature and Mean Diurnal Range correlated with germination rate. This suggests lower fitness due to maladaptation. Though we cannot exclude the possibility that reduced reproductive output could be an adaptive response to a poorer environment, given the short timescales involving a handful of generations this seems unlikely.

Based on our inference that populations with low c-RONA are more locally adapted, we then performed a comparison between c-RONA and the Shapley Index based on neutral diversity.

We find that populations with the highest inferred conservation value (highest Shapley score for neutral loci) were also those with the greatest deviation from optimum allele frequencies (highest c-RONA) (Table 2, Figure 3). This implies that it may be inappropriate to use the Shapley Index (and by extension, other similar metrics) based solely on neutral diversity for conservation prioritization, since this strategy would inadvertently favour poorly adapted populations that display a high degree of unique variation – in the case of dwarf birch, this is most likely due to genetic drift. Instead, we propose a conservation framework where populations with a low c-RONA and high Shapley Index based instead on adaptive diversity are prioritized. This would maximize both local adaptation and adaptive diversity, supporting future adaptive potential (Table S9).

To illustrate a possible application for this prioritization framework, we sought to identify putative dwarf birch donor populations that possess adaptive alleles at frequencies that would display reduced c-RONA in a recipient population (Figures 4, S7). We chose to demonstrate our approach using a current climate reference, as it could be considered more conservative, though we note that planning for future climate may have a better chance of long-term success. In this example, our hypothetical AGF strategy involves a substantial translocation of genotypes, particularly from the centre of the range towards the periphery. Whilst controversial, AGF may be advantageous, as it can introduce or increase the frequency of preadapted alleles to allow more rapid adaptation to track changing climate, alleviate inbreeding depression or increase adaptive potential (Frankham, 2015; Prober et al., 2015); and in the process provide a demographic safeguard by augmenting population size (Hodgins and Moore, 2016). In practice, implementation of AGF is likely to take the form of composite provenancing, whereby genetic material from a combination of source populations is used (Breed et al., 2013; Hodgins and Moore, 2016). This may seek to target adaptive diversity across multiple important environmental variables from across the species range, sometimes

irrespective of the distance to the source population and the 'local is best' paradigm (Boshier et al., 2015; Havens et al., 2015; Jones, 2013).

Our suggested approach has some limitations: RADseq only identifies variation in a subset of the genome (Lowry et al., 2016) possibly missing important adaptive loci (Harrisson et al., 2014). This concern may be addressed in future by whole genome population sequencing, and a better understanding of the limiting returns from typing more adaptive loci (for example Ahrens *et al.* 2018). Second, our approach does not explicitly account for phenotypic plasticity or the adaptive input from new mutations (Chevin and Lande, 2011). More generally, we caution against interpreting the statistical association between the allele frequency and the bioclimatic variates (for example, MDR) as a demonstration that the allele in question is linked to a quantitative trait locus with adaptive variation for that variable. Rather, the causal environmental variable may be unmeasured, but closely correlated with MDR. Finally, we highlight that, in our study area, the climate has been changing, albeit slowly, for several millennia, with the rate of climate change increasing more recently (Wang et al., 2014). Therefore, the clines identified here could represent adaptation to the environment of the recent past, rather than the present, and therefore may underestimate the current ecological risk. Negative density dependence could also obscure the effects of abiotic gradients. In the future, methods to accommodate change in the relative importance of environmental variables through time (Clark *et al.* 2014) and non-linear associations (Fitzpatrick and Keller 2015) are likely to advance our understanding and improve estimates of local adaptation in wild populations.

## Conclusions

Estimating the degree of maladaptation in populations as a criterion to inform selection of plant material for genetic rescue, composite provenancing or species reintroductions is currently the subject of considerable interest (Gibson et al., 2016; Leroy et al., 2018), and this is likely to

increase in the context of environmental change (Aitken and Bemmels, 2016). Here we present an approach to permit rapid assessment of local adaptation and future adaptive potential in wild populations. Importantly, the estimation of maladaptation presents a testable hypothesis; specifically, that if an AGF programme translocated individuals to a site where they are expected to display reduced c-RONA, the response of measurable fitness proxies such as catkin production should be positive. In dwarf birch, AGF would have to be combined with other management interventions focused on mitigating grazing pressure and burning to support natural regeneration, with the aim that larger populations eventually support 'natural' gene flow. Similarly, AGF need not entail translocation of genetic material to an existing recipient population in the first instance. Initially individuals of different provenance (and known allele frequencies) could be translocated to trial locations and subsequent fitness assessments would enable validation of the predicted adaptive potential. Conservationists and practitioners would then be in a better position to manage and, where appropriate, facilitate adaptation.

## Data Archiving Statement

1. Illumina read data from RADseq libraries has been uploaded to the European Nucleotide Archive project PRJEB26807, sample accessions ERS2598190- ERS2598376

2. Species records are available directly from the NBN Gateway [https://data.nbn.org.uk/].

3. Climate data are available from http://www.worldclim.org/

## Literature Cited

Abebe, T.D., Naz, A.A., Léon, J., 2015. Landscape genomics reveal signatures of local adaptation in barley (Hordeum vulgare L.). Front. Plant Sci. 6. https://doi.org/10.3389/fpls.2015.00813

Ahrens, C.W., Rymer, P.D., Stow, A., Bragg, J., Dillon, S., Umbers, K.D.L., Dudaniec, R.Y., 2018. The search for loci under selection: trends, biases and progress. Mol. Ecol. 27, 1342–1356. https://doi.org/10.1111/mec.14549

Aitken, S.N., Bemmels, J.B., 2016. Time to get moving: Assisted gene flow of forest trees. Evol. Appl. 9, 271–290. https://doi.org/10.1111/eva.12293

Aitken, S.N., Whitlock, M.C., 2013. Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. Annu. Rev. Ecol. Evol. Syst. 44, 367–88. https://doi.org/DOI 10.1146/annurev-ecolsys-110512-135747

Aitken, S.N., Yeaman, S., Holliday, J.A., Wang, T., Curtis-McLane, S., 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. Evol. Appl. 1, 95–111. https://doi.org/10.1111/j.1752-4571.2007.00013.x

Alberto, F.J., Aitken, S.N., Alía, R., González-Martínez, S.C., Hänninen, H., Kremer, A., Lefèvre, F., Lenormand, T., Yeaman, S., Whetten, R., Savolainen, O., 2013. Potential for evolutionary responses to climate change - evidence from tree populations. Glob. Chang. Biol. 19, 1645–1661. https://doi.org/10.1111/gcb.12181

Alsos, I.G., Spjelkavik, S., Engelskjøn, T., 2003. Seed bank size and composition of Betula nana, Vaccinium uliginosum, and Campanula rotundifolia habitats in Svalbard and northern Norway. Can. J. Bot. 81, 220–231. https://doi.org/10.1139/b03-018

Alsos, I.G., Spjelkavik, S., Engelskjøn, T., 2003. Seed bank size and composition of *Betula nana* , *Vaccinium uliginosum* , and *Campanula rotundifolia* habitats in Svalbard and northern Norway. Can. J. Bot. 81, 220–231. https://doi.org/10.1139/b03-018

Aston, D., 1984. *Betula nana* L., a note on its status in the United Kingdom. Proc. R. Soc. Edinburgh Sect. B Biol. Sci. 85, 43–47.

Bay, R.A., Rose, N.H., Logan, C.A., Palumbi, S.R., 2017. Genomic models predict successful coral adaptation if future ocean warming rates are reduced. Sci. Adv. 3, e1701413. https://doi.org/10.1126/sciadv.1701413

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., Courchamp, F., 2012. Impacts of Climate Change on the Future of Biodiversity. Ecol. Lett. 15, 365–377. https://doi.org/10.1111/j.1461-0248.2011.01736.x.Impacts

Bierne, N., Roze, D., Welch, J.J., 2013. Pervasive selection or is it ...? why are F ST outliers sometimes so frequent ? Mol. Ecol. 33, 2061–2064. https://doi.org/10.1111/mec.12241

Blair, L.M., Granka, J.M., Feldman, M.W., 2014. On the stability of the Bayenv method in assessing human SNP-environment associations. Hum. Genomics 8, 1–13.

https://doi.org/10.1186/1479-7364-8-1

Bonin, A., Nicole, F., Pompanon, F., Miaud, C., Taberlet, P., 2007. Population adaptive index: A new method to help measure intraspecific genetic diversity and prioritize populations for conservation. Conserv. Biol. 21, 697–708. https://doi.org/10.1111/j.1523-1739.2007.00685.x

Borrell, J.S., Wang, N., Nichols, R.A., Buggs, R.J.A., 2018. Genetic diversity maintained among fragmented populations of a tree undergoing range contraction. Heredity (Edinb). 121, 304–318. https://doi.org/10.1038/s41437-018-0132-8

Boshier, D., Broadhurst, L., Cornelius, J., Gallo, L., Koskela, J., Loo, J., Petrokofsky, G., St Clair, B., 2015. Is local best? Examining the evidence for local adaptation in trees and its scale. Environ. Evid. 4, 1–10. https://doi.org/10.1186/s13750-015-0046-3

Breed, M.F., Stead, M.G., Ottewell, K.M., Gardner, M.G., Lowe, A.J., 2013. Which provenance and where? Seed sourcing strategies for revegetation in a changing environment. Conserv. Genet. 14, 1–10. https://doi.org/10.1007/s10592-012-0425-z

Bryant, D., Moulton, V., 2004. Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. Mol. Biol. Evol. 21, 255–265. https://doi.org/10.1093/molbev/msh018

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140. https://doi.org/10.1111/mec.12354

Chevin, L.M., Lande, R., 2011. Adaptation to marginal habitats by evolution of increased phenotypic plasticity. J. Evol. Biol. 24, 1462–1476. https://doi.org/10.1111/j.1420-

9101.2011.02279.x

Clark, J.S., Gelfand, A.E., Woodall, C.W., Zhu, K., 2014. More than the sum of the parts: Forest climate response from joint species distribution models. Ecol. Appl. 24, 990–999. https://doi.org/10.1890/13-1015.1

Collen, B., Turvey, S.T., Waterman, C., Meredith, H.M.R., Kuhn, T.S., Baillie, J.E.M., Isaac, N.J.B., 2011. Investing in evolutionary history: implementing a phylogenetic approach for mammal conservation. Philos. Trans. R. Soc. B Biol. Sci. 366, 2611–2622. https://doi.org/10.1098/rstb.2011.0109

Coop, G., Witonsky, D., Rienzo, A. Di, Pritchard, J.K., 2010. Using Environmental Correlations to Identify Loci Underlying Local Adaptation 1423, 1411–1423. https://doi.org/10.1534/genetics.110.114819

Corlett, R.T., Westcott, D.A., 2013. Will plant movements keep up with climate change? Trends Ecol. Evol. 28, 482–488. https://doi.org/10.1016/j.tree.2013.04.003

Davis, M.B., Shaw, R.G., Ettersonm, J.R., 2005. Evolutionary Responses to Changing Climate. Ecology 86, 1704–1714.

De Mita, S., Thuillet, A.C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., Vigouroux, Y., 2013. Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. Mol. Ecol. 22, 1383–1399. https://doi.org/10.1111/mec.12182

Eizaguirre, C., Baltazar-Soares, M., 2014. Evolutionary conservation-evaluating the adaptive potential of species. Evol. Appl. 7, 963–967. https://doi.org/10.1111/eva.12227

Elith, J., Leathwick, J., 2009. Species distribution models: ecological explanation and prediction

across space and time. Annu. Rev. Ecol. Evol. ... 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Engler, R., Randin, C.F., Thuiller, W., Dullinger, S., Zimmermann, N.E., Araújo, M.B., Pearman, P.B., Le Lay, G., Piedallu, C., Albert, C.H., Choler, P., Coldea, G., De Lamo, X., Dirnböck, T., Gégout, J.C., Gómez-García, D., Grytnes, J.A., Heegaard, E., Høistad, F., Nogués-Bravo, D., Normand, S., Puşcaş, M., Sebastià, M.T., Stanisci, A., Theurillat, J.P., Trivedi, M.R., Vittoz, P., Guisan, A., 2011. 21st century climate change threatens mountain flora unequally across Europe. Glob. Chang. Biol. 17, 2330–2341. https://doi.org/10.1111/j.1365-2486.2010.02393.x

Fitzpatrick, M.C., Keller, S.R., 2015. Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. Ecol. Lett. 18, 1–16. https://doi.org/10.1111/ele.12376

Flanagan, S.P., Forester, B.R., Latch, E.K., Aitken, S.N., Hoban, S., 2018. Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. Evol. Appl. 11, 1035–1052. https://doi.org/10.1111/eva.12569

Foll, M., Gaggiotti, O., 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics 180, 977–993. https://doi.org/10.1534/genetics.108.092221

Forester, B.R., Lasky, J.R., Wagner, H.H., Urban, D.L., 2018. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Mol. Ecol. 27, 2215–2233. https://doi.org/https://doi.org/10.1111/mec.14584

François, O., Martins, H., Caye, K., Schoville, S.D., 2016. Controlling false discoveries in genome scans for selection. Mol. Ecol. 25, 454–469. https://doi.org/10.1111/mec.13513

Frankham, R., 2015. Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. Mol. Ecol. 24, 2610–2618. https://doi.org/10.1111/mec.13139

Frankham, R., Ballou, J.D., Eldridge, M.D.B., Lacy, R.C., Ralls, K., Dudash, M.R., Fenster, C.B., 2011. Predicting the Probability of Outbreeding Depression. Conserv. Biol. 25, 465–475. https://doi.org/10.1111/j.1523-1739.2011.01662.x

Funk, W.C., Forester, B.R., Converse, S.J., Darst, C., Morey, S., 2019. Improving conservation policy with genomics: a guide to integrating adaptive potential into U . S . Endangered Species Act decisions for conservation practitioners and geneticists. Conserv. Genet. 20, 115–134. https://doi.org/10.1007/s10592-018-1096-1

Gent, P.R., Danabasoglu, G., Donner, L.J., Holland, M.M., Hunke, E.C., Jayne, S.R., Lawrence, D.M., Neale, R.B., Rasch, P.J., Vertenstein, M., Worley, P.H., Yang, Z.L., Zhang, M., 2011. The community climate system model version 4. J. Clim. 24, 4973–4991. https://doi.org/10.1175/2011JCLI4083.1

Gibson, A.L., Espeland, E.K., Wagner, V., Nelson, C.R., 2016. Can local adaptation research in plants inform selection of native plant materials? An analysis of experimental methodologies. Evol. Appl. 9, 1219–1228. https://doi.org/10.1111/eva.12379

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., A, T.D., I, A., X, A., L, F., R, R., Q, Z., Z, C., E, M., N, H., Gnirke, A., Rhind, N., Palma, F. di, W., B., Friedman, N., Regev, A., 2013. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883.Trinity

Gratani, L., 2014. Plant Phenotypic Plasticity in Response to Environmental Factors. Adv. Bot. 2014, 1–17. https://doi.org/10.1155/2014/208747

Gumbs, R., Gray, C.L., Wearn, O.R., Owen, N.R., 2018. Tetrapods on the EDGE: Overcoming data limitations to identify phylogenetic conservation priorities. PLoS One 13, 1–19. https://doi.org/10.1371/journal.pone.0194680

Günther, T., Coop, G., 2013. Robust identification of local adaptation from allele frequencies. Genetics 195, 205–220. https://doi.org/10.1534/genetics.113.152462

Haake, C.-J., Kashiwada, A., Su, F.E., 2007. The Shapley value of phylogenetic trees. J. Math. Biol. 56, 479–497. https://doi.org/10.1007/s00285-007-0126-2

Hällfors, M.H., Liao, J., Dzurisin, J.D.K., Grundel, R., Hyvärinen, M., Towle, K., Wu, G.C., Hellmann, J.J., 2015. Addressing potential local adaptation in species distribution models: implications for conservation under climate change. Ecol. Appl. 15–0926.1. https://doi.org/10.1890/15-0926.1

Hampe, A., Petit, R.J., 2005. Conserving biodiversity under climate change: The rear edge matters. Ecol. Lett. 8, 461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

Harrisson, K.A., Pavlova, A., Telonis-Scott, M., Sunnucks, P., 2014. Using genomics to characterize evolutionary potential for conservation of wild populations. Evol. Appl. 7, 1008–1025. https://doi.org/10.1111/eva.12149

Havens, K., Vitt, P., Still, S., Kramer, A.T., Fant, J.B., Schatz, K., 2015. Seed Sourcing for Restoration in an Era of Climate Change. Nat. Areas J. 35, 122–133. https://doi.org/10.3375/043.035.0116

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. https://doi.org/10.1002/joc.1276

Hijmans, R.J., Etten, J., 2012. raster: Geographic analysis and modeling with raster data.

Hijmans, R.J., Philips, S., Leathwick, J., Elith, J., 2011. Package "dismo."

Hodgins, K.A., Moore, J.L., 2016. Adapting to a warming world: Ecological restoration, climate change, and genomics. Am. J. Bot. 103, 590–592. https://doi.org/10.3732/ajb.1600049

Hoersch, B., Braun, G., Schmidt, U., 2002. Relation between landform and vegetation in alpine regions of Wallis, Switzerland. A multiscale remote sensing and GIS approach. Comput. Environ. Urban Syst. 26, 113–139. https://doi.org/10.1016/S0198-9715(01)00039-4

Hoffmann, A.A., Sgrò, C.M., Kristensen, T.N., 2017. Revisiting Adaptive Potential, Population Size, and Conservation. Trends Ecol. Evol. 32, 506–517. https://doi.org/10.1016/j.tree.2017.03.012

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267. https://doi.org/10.1093/molbev/msj030

IPCC, 2014a. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)].

IPCC, 2014b. Climate Change 2014 Synthesis Report Summary Chapter for Policymakers. Ipcc 31. https://doi.org/10.1017/CBO9781107415324

Isaac, N.J.B., Turvey, S.T., Collen, B., Waterman, C., Baillie, J.E.M., 2007. Mammals on the EDGE: Conservation priorities based on threat and phylogeny. PLoS One 2. https://doi.org/10.1371/journal.pone.0000296

Jeffery, H., 1961. The Theory of Probability, 6th ed., 3rd. ed. ed, Oxford University Press,

Oxford. Classic Texts in the Physical Sciences . Oxford Univ. Press, Oxford.

Jetz, W., Thomas, G.H., Joy, J.B., Redding, D.W., Hartmann, K., Mooers, A.O., 2014. Global Distribution and Conservation of Evolutionary Distinctness in Birds. Curr. Biol. 24, 919–930. https://doi.org/10.1016/j.cub.2014.03.011

Jin, Y.K., Lundgren, P., Lutz, A., Raina, J.-B., Howells, E.J., Paley, A.S., Willis, B.L., van Oppen, M.J.H., 2016. Genetic markers for antioxidant capacity in a reef-building coral. Sci. Adv. 2, e1500842–e1500842. https://doi.org/10.1126/sciadv.1500842

Jones, T.A., 2013. When local isn't best. Evol. Appl. 6, 1109–1118. https://doi.org/10.1111/eva.12090

Kardos, M., Shafer, A.B.A., 2018. The Peril of Gene-Targeted Conservation. Trends Ecol. Evol. 33, 827–839. https://doi.org/10.1016/j.tree.2018.08.011

Kawecki, T.J., Ebert, D., 2004. Conceptual issues in local adaptation. Ecol. Lett. 7, 1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

Keller, S.R., Levsen, N., Olson, M.S., Tiffin, P., 2012. Local adaptation in the flowering-time gene network of balsam poplar, populus balsamifera L. Mol. Biol. Evol. 29, 3143–3152. https://doi.org/10.1093/molbev/mss121

Koskela, J., Lefèvre, F., Schueler, S., Kraigher, H., Olrik, D.C., Hubert, J., Longauer, R., Bozzano, M., Yrjänä, L., Alizoti, P., Rotach, P., Vietto, L., Bordács, S., Myking, T., Eysteinsson, T., Souvannavong, O., Fady, B., De Cuyper, B., Heinze, B., von Wühlisch, G., Ducousso, A., Ditlevsen, B., 2013. Translating conservation genetics into management: Pan-European minimum requirements for dynamic conservation units of forest tree genetic diversity. Biol. Conserv. https://doi.org/10.1016/j.biocon.2012.07.023

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat Meth 9, 357–359.

Legendre, P., Legendre, L., 2012. Numerical Ecology, 3rd editio. ed. Elsevier, Amsterdam.

Leimu, R., Fischer, M., 2008. A meta-analysis of local adaptation in plants. PLoS One 3, 1–8. https://doi.org/10.1371/journal.pone.0004010

Leroy, G., Carroll, E.L., Bruford, M.W., DeWoody, J.A., Strand, A., Waits, L., Wang, J., 2018. Next-generation metrics for monitoring genetic erosion within populations of conservation concern. Evol. Appl. 11, 1066–1083. https://doi.org/10.1111/eva.12564

Liu, C., Newell, G., White, M., 2016. On the selection of thresholds for predicting species occurrence with presence-only data. Ecol. Evol. 6, 337–348. https://doi.org/10.1002/ece3.1878

Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B., Ackerly, D.D., 2009. The velocity of climate change. Nature 462, 1052–1055. https://doi.org/10.1038/nature08649

Lowry, D.B., Hoban, S., Kelley, J.L., Lotterhos, K.E., Reed, L.K., Antolin, M.F., Storfer, A., 2016. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Mol. Ecol. Resour. 17, 142–152. https://doi.org/10.1111/1755-0998.12635

Maguire, K.C., Nieto-Lugilde, D., Fitzpatrick, M.C., Williams, J.W., Blois, J.L., 2015. Modeling Species and Community Responses to Past, Present, and Future Episodes of Climatic and Ecological Change. Annu. Rev. Ecol. Evol. Syst. 46, 343–368.

Manthey, J.D., Moyle, R.G., 2015. Isolation by environment in white-breasted nuthatches (Sitta carolinensis) of the Madrean Archipelago sky islands: A landscape genomics approach.

Mol. Ecol. 24, 3628–3638. https://doi.org/10.1111/mec.13258

McKenney, D., Pedlar, J., Lawrance, K., Campbell, K., Hutchinson, M., 2007. Potential impacts of climate change on the distribution of North American trees. Bioscience 57, 939–948. https://doi.org/10.1641/b571106

Meier, E.S., Lischke, H., Schmatz, D.R., Zimmermann, N.E., 2012. Climate, competition and connectivity affect future migration and ranges of European trees. Glob. Ecol. Biogeogr. 21, 164–178. https://doi.org/10.1111/j.1466-8238.2011.00669.x

Nadeau, S., Meirmans, P.G., Aitken, S.N., Ritland, K., Isabel, N., 2016. The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. Ecol. Evol. 6, 8649–8664. https://doi.org/10.1002/ece3.2550

Nicotra, A.B., Atkin, O.K., Bonser, S.P., Davidson, A.M., Finnegan, E.J., Mathesius, U., Poot, P., Purugganan, M.D., Richards, C.L., Valladares, F., van Kleunen, M., 2010. Plant phenotypic plasticity in a changing climate. Trends Plant Sci. 15, 684–692. https://doi.org/10.1016/j.tplants.2010.09.008

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.M., Wagner, H., 2019. Package "vegan": Community ecology package.

Olson, M.S., Levsen, N., Soolanayakanahally, R.Y., Guy, R.D., Schroeder, W.R., Keller, S.R., Tiffin, P., 2013. The adaptive potential of Populus balsamifera L. to phenology requirements in a warmer global climate. Mol. Ecol. 22, 1214–1230. https://doi.org/10.1111/mec.12067

Pardo-Diaz, C., Salazar, C., Jiggins, C.D., 2015. Towards the identification of the loci of adaptive

evolution. Methods Ecol. Evol. 6, 445–464. https://doi.org/10.1111/2041-210X.12324

Parolo, G., Rossi, G., Ferrarini, A., 2008. Assessing ecological responses to environmental change using statistical models. J. Appl. Ecol. 45, 1410–1418. https://doi.org/10.1111/j.1365-2664.2007.0

Pavlova, A., Beheregaray, L.B., Coleman, R., Gilligan, D., Harrisson, K.A., Ingram, B.A., Kearns, J., Lamb, A.M., Lintermans, M., Lyon, J., Nguyen, T.T.T., Sasaki, M., Tonkin, Z., Yen, J.D.L., Sunnucks, P., 2017. Severe consequences of habitat fragmentation on genetic diversity of an endangered Australian freshwater fish: A call for assisted gene flow. Evol. Appl. 10, 531–550. https://doi.org/10.1111/eva.12484

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pina-Martins, F., Baptista, J., Pappas Jr G. & Paulo, O.S., 2018. New insights on adaptation and population structure of cork oak using genotyping by sequencing 0–3. https://doi.org/10.1101/263160

Polechová, J., 2018. Is the sky the limit? On the expansion threshold of a species' range. PLoS Biol. 16, 1–18. https://doi.org/10.1371/journal.pbio.2005372

Polechová, J., Barton, N.H., 2015. Limits to adaptation along environmental gradients. Proc. Natl. Acad. Sci. 112, 6401–6406. https://doi.org/10.1073/pnas.1421515112

Prober, S.M., Byrne, M., McLean, E.H., Steane, D.A., Potts, B.M., Vaillancourt, R.E., Stock, W.D., 2015. Climate-adjusted provenancing: a strategy for climate-resilient ecological restoration. Front. Ecol. Evol. 3, 1–5. https://doi.org/10.3389/fevo.2015.00065

R Development Core Team, 2014. R: A language and environment for statistical computing. Found. Stat. Comput. Vienna, Austria.

Razgour, O., Forester, B., Taggart, J.B., Bekaert, M., Juste, J., Ibáñez, C., 2019. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections 116, 6–11. https://doi.org/10.1073/pnas.1820663116

Reed, D.H., Frankham, R., 2003. Correlation between Fitness and Genetic Diversity\rCorrelación entre Adaptabilidad y Diversidad Genética. Conserv. Biol. 17, 230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Rellstab, C., Gugerli, F., Eckert, A.J., Hancock, A.M., Holderegger, R., 2015. A practical guide to environmental association analysis in landscape genomics. Mol. Ecol. 24, 4348–4370. https://doi.org/10.1111/mec.13322

Rellstab, C., Zoller, S., Walthert, L., Lesur, I., Pluess, A.R., Graf, R.R., Bodénès, C., Sperisen, C., Kremer, A., Gugerli, F., Bodï¿½nï¿½s, C., Sperisen, C., Kremer, A., Gugerli, F., 2016. Signatures of local adaptation in candidate genes of oaks (Quercus spp.) with respect to present and future climatic conditions. Mol. Ecol. 25, 5907–5924. https://doi.org/10.1111/mec.13889

Schweizer, R.M., Vonholdt, B.M., Harrigan, R., Wayne, R.K., 2016. Genetic subdivision and candidate genes under selection in North American grey wolves 380–402. https://doi.org/10.1111/mec.13364

Searcy, C.A., Shaffer, H.B., 2016. Do Ecological Niche Models Accurately Identify Climatic Determinants of Species Ranges? Am. Nat. 187:4, 423–435.

Urban, M.C., 2015. Accelerating extinction risk from climate change. Science (80-. ). 348.

Volkmann, L., Martyn, I., Moulton, V., Spillner, A., Mooers, A.O., 2014. Prioritizing populations for conservation using phylogenetic networks. PLoS One 9. https://doi.org/10.1371/journal.pone.0088945

Wang, N., Borrell, J.S., Bodles, W.J. a, Kuttapitiya, A., Nichols, R. a., Buggs, R.J. a, 2014. Molecular footprints of the Holocene retreat of dwarf birch in Britain. Mol. Ecol. 23, 2771–2782. https://doi.org/10.1111/mec.12768

Wang, N., Thomson, M., Bodles, W.J.A. a, Crawford, R.M.M.M., Hunt, H. V., Featherstone, A.W., Pellicer, J., Buggs, R.J.A. a, 2013. Genome sequence of dwarf birch (Betula nana) and cross-species RAD markers. Mol. Ecol. 22, 3098–3111. https://doi.org/10.1111/mec.12131

Wang, T., Wang, Z., Xia, F., Su, Y., 2016. Local adaptation to temperature and precipitation in naturally fragmented populations of Cephalotaxus oliveri, an endangered conifer endemic to China. Sci. Rep. 6, 1–12. https://doi.org/10.1038/srep25031

Weeks, A.R., Stoklosa, J., Hoffmann, A.A., 2016. Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: The case of Australian mammals. Front. Zool. 13, 1–9. https://doi.org/10.1186/s12983-016-0163-z

Williams, S.E., Moritz, C., Shoo, L.P., Isaac, J.L., Hoffmann, A. a, Langham, G., 2008. Towards an Integrated Framework for Assessing the Vulnerability of Species to Climate Change. PLoS Biol 6, e325. https://doi.org/10.1371/journal.pbio.0060325

Zhu, K., Woodall, C.W., Clark, J.S., 2012. Failure to migrate: Lack of tree range expansion in response to climate change. Glob. Chang. Biol. 18, 1042–1052. https://doi.org/10.1111/j.1365-2486.2011.02571.x

Zohren, J., 2016. Introgression in Betula Species of Different Ploidy Levels and the Analysis of the Betula nana Genome. Queen Mary University of London.

Zohren, J., Wang, N., Kardailsky, I., Borrell, J.S., Joecker, A., Nichols, R.A., Buggs, R.J.A., 2016. Unidirectional diploid-tetraploid introgression among British birch trees with shifting ranges shown by restriction site-associated markers. Mol. Ecol. 25, 2413–2426. https://doi.org/10.1111/mec.13644



# Tables

**Table 1.** Contribution of retained environmental variables to the dwarf birch environmental niche model (ENM), and the number of environmentally associated loci detected.

| Variable | Description | Correlated Variables[1] | ENM percent contribution[2] | GEA Loci | GEA Loci (inc. cor.)[3] |
|---|---|---|---|---|---|
| AMTemp | Annual Mean Temperature | MTColdQ, MTColdM | 34.9 | 17 | 64 |
| MTWarmM | Max Temperature of Warmest Month | MTWarmQ | 22.1 | 2 | 6 |
| MDR | Mean Diurnal Range | - | 14.8 | 71 | 71 |
| ISO | Isothermality | - | 14.6 | 11 | 11 |
| APrec | Annual Precipitation | PColdQ, PWetM, PSeason, PWetQ, | 7.3 | 2 | 21 |

| | | | | | |
|---|---|---|---|---|---|
| | | PWarmQ, PDryM, PDryQ | | | |
| Slope | Slope | - | 2.8 | 7 | 7 |
| MTDryQ | Mean Temperature of Driest Quarter | - | 1.6 | 7 | 7 |
| TS | Temperature Seasonality | ATempR | 1.4 | 1 | 3 |
| MTWetQ | Mean Temperature of Wettest Quarter | - | 0.3 | 7 | 7 |
| Aspect | Aspect | - | 0.2 | 4 | 4 |

[1] Correlated variables include Mean Temperature of the Coldest Quarter (MTColdQ); Minimum Temperature of the Coldest Month (MTColdQ); Mean Temperature of Warmest Quarter (MTWarmQ); Precipitation of Coldest Quarter (PColdQ); Precipitation of Wettest Month (PWetM); Precipitation Seasonality (Pseason); Precipitation of Wettest Quarter (PWetQ); Precipitation of the Warmest Quarter (PWarmQ); Precipitation of Driest Month (PDryM); Precipitation of Driest Quarter (PDryQ); Annual Temperature Range (ATempR).

[2] Percentage contribution is calculated as the increase in regularized gain added to the contribution of the corresponding variable over each iteration of the model.

[3] Total number of SNPs associated with both the retained variable, as well as related highly correlated variables that were excluded from the ENM analysis.

**Table 2.** Summary information for 29 dwarf birch populations, including the number of genotyped and phenotyped individuals, habitat suitability (HSI).

| Location | Pop. | Lat. | Long | Elev. (m) | Genotyped | Phenotyped | HSI | c-PONA | $Shapley_{NEUTRA}$ |
|---|---|---|---|---|---|---|---|---|---|
| Ben Loyal | BL | 58.4 | -4.4 | 300 | 6 | 30 | 0.39 | 0.194 | 0.011 |
| Meall Odhar | MO | 58.16 | -4.42 | 404 | 6 | 29 | 0.45 | 0.168 | 0.006 |
| Beinn Enaiglair | BE | 57.79 | -5.01 | 480 | 5 | 27 | 0.37 | 0.479 | 0.01 |
| Luichart | LH | 57.72 | -4.9 | 268 | 6 | 29 | 0.54 | 0.131 | 0.008 |
| Ben Wyvis W | BW | 57.65 | -4.6 | 482 | 5 | 30 | 0.77 | 0.149 | 0.01 |
| Ben Wyvis E* | DG | 57.65 | -4.56 | 472 | - | 21 | 0.75 | - | - |
| Loch Meig | ME | 57.53 | -4.8 | 450 | 6 | 26 | 0.57 | 0.128 | 0.005 |
| Glen Cannich | GC | 57.34 | -4.86 | 455 | 6 | 31 | 0.51 | 0.045 | 0.027 |
| Faskanyle* | FS | 57.33 | -4.85 | 486 | - | 17 | 0.66 | - | - |
| Dundreggan | DE | 57.23 | -4.75 | 448 | 6 | 30 | 0.81 | 0.174 | 0.009 |
| An Suidhe | AS | 57.22 | -4.81 | 661 | 2 | 17 | 0.77 | 0.219 | 0.119 |
| Beinn Bhreac | BB | 57.21 | -4.82 | 500 | 6 | 33 | 0.66 | 0.366 | 0.008 |
| Portclair | PC | 57.2 | -4.64 | 478 | 6 | 38 | 0.54 | 0.081 | 0.008 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| River Avon | AV | 57.14 | -3.49 | 549 | 6 | 28 | 0.59 | 0.306 | 0.01 |
| Monadhliaths | MD | 57.06 | -4.31 | 712 | 6 | 6 | 0.49 | 0.222 | 0.01 |
| Meall an tsluagain | SL | 57.05 | -3.45 | 633 | 6 | 31 | 0.59 | 0.085 | 0.035 |
| Loch Muick E | MU1 | 56.92 | -3.2 | 492 | 6 | 31 | 0.17 | 0.223 | 0.006 |
| Loch Muick W | MU2 | 56.92 | -3.21 | 517 | 6 | 16 | 0.1 | 0.218 | 0.008 |
| Loch Laggan | LG | 56.89 | -4.54 | 364 | 6 | 33 | 0.35 | 0.064 | 0.007 |
| Loch Loch | LL | 56.85 | -3.65 | 673 | 6 | 32 | 0.57 | 0.106 | 0.005 |
| Ben Gullabin | BG | 56.84 | -3.47 | 594 | 1 | 7† | 0.58 | 0.194 | 0.422 |
| Loch Rannoch | LR | 56.76 | -4.42 | 499 | 6 | 28 | 0.23 | 0.097 | 0.008 |
| Rannoch West | RW | 56.65 | -4.79 | 306 | 6 | 32 | 0.61 | 0.218 | 0.007 |
| Rannoch Moor B | RB | 56.6 | -4.74 | 304 | 6 | 10 | 0.51 | 0.169 | 0.008 |
| Rannoch Moor A* | RA | 56.6 | -4.74 | 295 | - | 27 | 0.51 | - | - |
| Lennox | LX | 55.97 | -4.28 | 164 | 2 | 10 | 0 | 0.241 | 0.102 |
| Emblehope† | EM | 55.24 | -2.48 | 448 | 1 | 1† | 0.06 | 0.254 | 0.155 |
| Spadeadam† | SA | 55.05 | -2.57 | 275 | 1 | 1† | 0.01 | 0.321 | 0.35 |
| Teesdale† | TD | 54.65 | -2.28 | 499 | 2 | 2† | 0.06 | 0.291 | 0.133 |

*Populations not submitted for genetic analysis, but are considered in the comparison of HSI and reproductive output.

†Populations were exhaustively sampled.

# Figure Legends

**Figure 1.** A) Environmental niche model of dwarf birch habitat suitability (HSI) under current environmental conditions, black points indicate species distribution records and red points indicate sampled locations included in this study. B) Regression of phenotypic fitness traits against the derived habitat suitability index. C) dwarf birch habitat suitability index projections under future climate scenarios.

**Figure 2.** Schematic diagram of current and future risk of non-adaptedness (c-RONA and f-RONA), presented on a genotype-environment association (GEA) plot; where genotypes are BLUP estimates of population polygenic allele frequency for 17 loci and the environmental predictor is Annual Mean Temperature. c/f-RONA is the average change in allele frequency

required to match our estimated optimum for current environmental conditions. Where RONA is large, we show two possible adaptation strategies; i) Assisted migration indicates the change in environmental conditions required for a population to match a genotype-environment optimum. This could take the form of a translocation of individuals to a location with a more suitable climate (e.g. a higher elevation). ii) Assisted Gene Flow (which in small populations is equivalent to Genetic Rescue) proposes movement of genetic material from a donor population with allele frequencies predicted to be better suited to the environmental conditions at the focal population. We show that the allele frequency change is likely to be larger under an example future climate scenario of 1°C warming. Blue and red bands indicate suitable candidate donor populations for assisted gene flow under current and future scenarios respectively.

**Figure 3.** A) Barplot of Shapley index for neutral and adaptive loci across UK *B. nana* populations, ordered by latitude with northernmost populations to the left. Inset plots show B) the relationship between the log transformed Shapley Index and C) the current risk of non-adaptedness (c-RONA) for neutral and adaptive loci respectively.

**Figure 4.** Hypothetical plots of assisted gene flow (AGF) for dwarf birch in the UK. Arrows denote movement from donor to recipient populations (red circles). Blue populations report an allele frequency close to predicted optimums, thus introduction of novel diversity does not decrease c-RONA and is not required. Base maps show A) Annual Mean Temperature (AMTemp) and B) Mean Diurnal Range (MDR) environmental variables.

**Figure 5.** A) The relationship between c-RONA (for AMTemp) and mean population catkin count. B) Correlation between the number of loci identified in genotype-environment analyses, for each environmental variable, and the corresponding percentage contribution of that variable to the environmental niche model.

917

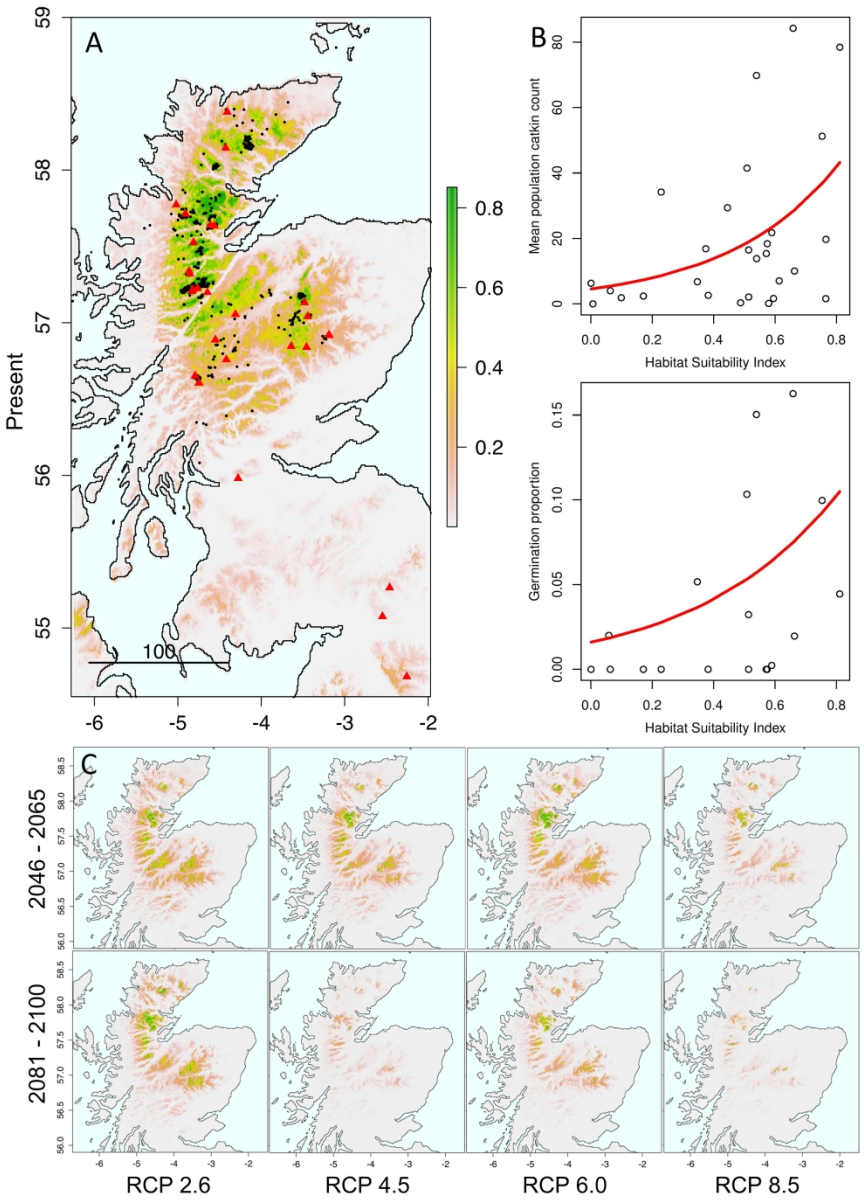


Figure 1. A) Environmental niche model of dwarf birch habitat suitability (HSI) under current environmental conditions, black points indicate species distribution records and red points indicate sampled locations included in this study. B) Regression of phenotypic fitness traits against the derived habitat suitability index. C) dwarf birch habitat suitability index projections under future climate scenarios.

201x275mm (600 x 600 DPI)